

## Implementasi *Data Analytic* Dalam Upaya Peningkatan Penjualan Properti Sebesar 10% Di NYC Amerika Serikat

**Fasha I. Husna**

Teknik Industri, Universitas Pembangunan Nasional “Veteran” Jawa Timur  
Jl. Rungkut Madya, Surabaya 60294  
e-mail : [21032010221@student.upnjatim.ac.id](mailto:21032010221@student.upnjatim.ac.id)

**Tranggono**

Teknik Industri, Universitas Pembangunan Nasional “Veteran” Jawa Timur  
Jl. Rungkut Madya, Surabaya 60294  
e-mail : [tranggono.ti@upnjatim.ac.id](mailto:tranggono.ti@upnjatim.ac.id)

**Abstract.** *The property industry is an example of an industry that is expected to continue to grow. Problems related to obstacles that occur in property transactions are the decline in the level of property sales in an area. This can happen due to several causes such as infrastructure gaps in each region, lack of promotion, etc. This research will answer the question of how to increase property sales by 10% every year in the New York City area of the United States. The dataset used in this research was taken via the Kaggle website and comes from information on the number of property sales in 5 boroughs in NYC from 2016-2017. It is hoped that the conclusions from the related journals will provide a recommendation model to increase property sales by 10% in the coming year in NYC.*

**Keywords :** *Property Industry, Sale, NYC*

**Abstrak.** Perindustrian properti merupakan contoh industri yang diperkirakan akan terus berkembang. Hal-hal terkait kendala yang terjadi dalam transaksi properti adalah menurunnya tingkat penjualan properti di suatu daerah. hal tersebut bisa terjadi dikarenakan beberapa penyebab seperti kesenjangan infrastruktur pada masing-masing daerah, kurangnya promosi, dsb. Dalam penelitian ini akan menjawab pertanyaan bagaimana cara meningkatkan penjualan *property* sebesar 10% setiap tahunnya di wilayah New York City Amerika Serikat. Dataset yang dipakai dari penelitian ini diambil melalui *website kaggle* dan berasal dari informasi jumlah penjualan properti 5 wilayah di NYC dari tahun 2016-2017. Kesimpulan dari jurnal terkait diharapkan akan memperoleh model rekomendasi untuk dapat menaikkan penjualan properti sebesar 10% pada tahun mendatang di NYC.

**Kata Kunci :** *Industri Properti, Penjualan, NYC*

### PENDAHULUAN

Perindustrian properti merupakan salah satu industri yang diperkirakan akan tetap ramai peminat. Salah satu kendala yang dihadapi dalam transaksi properti yakni kesulitan memperoleh data informasi terkait harga jual properti yang wajar. hal itu dapat terjadi dikarenakan beberapa penyebab seperti kesenjangan infrastruktur pada masing-masing daerah, kurangnya promosi, dsb. Pada zaman modern ini, biasanya data perihal properti sangat mungkin didapat dari internet lebih gampang. Data terkait umumnya tersebar pada website transaksi properti, situs milik pemerintah, atau bisa pada media digital. Tetapi, bukan seluruh

data yang tersedia di internet mampu menggambarkan harga properti di wilayah terkait. Memanfaatkan penggunaan inovasi big data, data-data terkait bisa didapat dan dikelompokkan dengan ukuran dan jumlah yang banyak berdasar melalui banyak sumber. Big data diartikan sebuah sistem yang mampu menyimpan banyak tipe data dengan ukuran yang besar dan kemampuan pengaturan data yang sangat cepat. Selain menyimpan, big data juga bisa menjalankan pengerjaan data dengan memanfaatkan fitur analisis big data. Pada masalah terkait, penerapan big data bisa menolong untuk menyimpan informasi properti, menjalankan informasi terkait berpatokan pada permintaan masyarakat yang memanfaatkannya, dan memberi kesimpulan proses yang sebaiknya didapatkan oleh *user*.

Dengan mendapatkan dataset penjualan properti di New York City Amerika Serikat, kami berupaya untuk mempelajari data tersebut dengan tujuan untuk identifikasi area populer, memahamai faktor-faktor yang mempengaruhi nilai properti, dan mengungkap tren tersembunyi yang dapat menjadi masukan bagi investasi masa depan. Dengan objektivitas yakni menganalisis kinerja penjualan setahun terakhir untuk memaksimalkan penjualan di tahun berikutnya dan memberikan rekomendasi.

Dalam penelitian ini akan menjawab pertanyaan bagaimana cara meningkatkan penjualan *property* sebesar 10% setiap tahunnya di wilayah New York City Amerika Serikat. Dataset yang akan dipakai dalam penelitian ini diambil melalui *website kaggle* dan berasal dari informasi jumlah penjualan properti 5 wilayah di NYC dari tahun 2016-2017. Kesimpulan dari jurnal terkait diharap akan memperoleh sistem model rekomendasi untuk dapat meningkatkan penjualan properti sebesar 10% pada tahun mendatang di NYC.

## TINJAUAN PUSTAKA

### A. Pengertian Statistika

Statistika diketahui secara luas sebagai salah satu disiplin keilmuan yang memahami teknik-teknik pengambilan kesimpulan suatu *problem* dengan memanfaatkan setengah informasi kuantitatif dari *problem* terkait. Setiap informasi disebut datum ( bahasa Latin artinya keterangan ). Karena pada dasarnya kita perlu banyak informasi sebagai bahan untuk menimbang pengambilan suatu kesimpulan, jadi dipakai kata jamak dari datum, yaitu data (keterangan-keterangan).

Statistika adalah bentuk dari angka yang dibentuk dalam tabel yang memberikan keterangan suatu peristiwa, untuk datanya didapat melalui kesimpulan analisis yang ditampilkan melalui model angka-angka. Statistik juga didefinisikan sekumpulan data yang dipergunakan untuk menjelaskan kendala dan menarik kesimpulan yang tepat. Statistik

mencakup pengelompokan data, penugasan sebuah data, analisis sebuah data, dan mencari hasil akhir berpatokan pada data yang tersedia. Data merupakan sekumpulan informasi yang bisa menjelaskan suatu keadaan bahan suatu *problem*. Sebuah data dapat dibidang mencakupi syarat bilamana obyektif yang diartikan jika data dikelompokkan harus tersedia gambaran keadaan *real*, relevan diartikan jika sebuah data yang dikelompokkan punya hubungan dengan sebuah kendala yang akan dianalisis, representative diartikan semisal data yang dikelompokkan dalam teknik sampling, sama dengan era, harus dapat mewakili dan menjelaskan keadaan populasinya, dapat dipercaya yakni definisinya jika sebuah data yang dikumpulkan didapat dari dasar yang valid (Luthfiana, 2020).

#### A. Pengertian Data Mining

Data mining punya definisi lain yakni *knowledge discovery* atau *pattern recognition* adalah pengertian yang dipakai dalam memperoleh informasi yang rahasia dari segerombolan data aturan yang sangat besar. maksud akhir data mining yakni untuk mencari, menggali, atau menabung informasi dari data atau informasi yang di punyai. proses data mining yaitu mencari pola yang bermakna dari pola dengan menerapkan teknik statistik dan matematika untuk mengurutkan sejumlah banyak data yang disimpan dalam tempat penyimpanan, yang berbasis IT nalar pola serta metode statistik dan matematis.

Menurut Larose, data mining dipecah dalam beberapa grup berpatokan pada tugas tugas yang mampu dijalankan, yaitu Deskripsi, Untuk menemukan metode yang menjelaskan pola serta kecenderungan yang ada dalam informasi. Penjelasan dari metode dan kecenderungan yang kerap membagikan mungkin penjelasan kemungkinan pola atau kecenderungan. Perkiraan diprediksi nyaris sepadan dengan klasifikasi, terkecuali variabel sasaran ditaksir lebih kearah numeric dibanding kearah jenis. Model dibentuk memakai rekaman komplit yang sediakan nilai melalui variabel sasaran selaku nilai prediksi. Prediksi setengah mirip terkait klasifikasi dan perkiraan, pengecualian jika melalui prediksi nilai dari kesimpulan terakhir akan tersedia di masa depan. Klasifikasi, Pada proses klasifikasi, ada tujuan variabel kategori sebagai pandangan, pengklusteran penghasilan bisa dipecah dalam tiga jenis, yakni penghasilan tinggi, menengah dan rendah. Pengklusteran, Pengklusteran ialah pengelompokkan *record*, pengamatan ataupun mencermati serta membentuk kelas objek yang mempunyai kesamaan. Kluster merupakan segerombol *record* yang mempunyai kesamaan dengan yang yang lain serta mempunyai perbedaan dengan *record* dalam kluster lain. Tahapan pengklusteran berupaya buat melaksanakan pengelompokkan terkait totalitas informasi jadi grup yang mempunyai kesamaan (homogen), yang mana kesamaan record pada satu grup

hendak bernilai optimal, sebaliknya kesamaan terkait *record* pada kelompok berbeda hendak bernilai minimum. Asosiasi, pekerjaan asosiasi pada data mining merupakan menciptakan atribut yang timbul dalam sesuatu waktu. Pada penerapan dunia bisnis lebih universal diucap penelitian keranjang belanja (Rizal, 2022).

### B. Pengertian Data Analyst

Permasalahan untuk menjadi data analyst pada dunia kerja yakni permintaan agar mengolah data pada macam-macam sumber, namun data yang didapatkan dikelompokkan pada macam-macam media berpatokan pada tipe data dan dasarnya. Dengan besarnya data yang dipunyai pegawai, susah untuk mengelompokkan kualitas data. Ketidaksiapan untuk memantau data secara nyata memiliki sekelompok dampak buruk yang setipeterhadap performa keputusan yang diciptakan berpatokan data yang disediakan maupun yang dikenal sebagai pengambilan keputusan berbasis data.

Menurut laporan yang dikeluarkan oleh Forbes pada tahun 2020, lowongan kerja untuk pekerjaan data analyst diprediksi bakal bertambah dari 36.000 bertambah ke 2.720.000. Meninjau angka terkait, pekerjaan ini ramai digemari. Tetapi, data analyst yang terpicat masih sangat minim, ditinjau dari melambungnya minat. Di Indonesia, masih kurang SDM yang punya keahlian data analyst dan pengalaman calon pekerja.

Analisis data / Data Analyst merupakan proses mencari, mengelola, manipulasi, dan memodelkan data dengan tujuan mencari informasi yang bermanfaat, menarik kesimpulan, dan mendukung pengambilan keputusan. Analisis data merupakan tahap disesuaikannya besaran data dan pengorganisasian data ke dalam pola, jenis, dan unit deskriptif dasar. Analisis data diartikan tahap mencari tema yang tersembunyi oleh data, merinci usaha formal agar menciptakan hipotesis dan mencoba mendukung tema dan hipotesis tersebut. Analisis data diartikan tahapan mengatur dan mengelompokkan data ke dalam pola, jenis, dan unit deskriptif dasar agar menemukan tema yang direkomendasikan data dan melakukan pengembangan hipotesis kerja (Susanto, 2023).

### C. Industri Properti

Perindustrian properti di Indonesia merupakan contoh industri yang bakal senantiasa banyak peminat. Semakin bertambahnya waktu masyarakat di Indonesia bakal terus melambung dan tanah atau lokasi tinggal yang tersedia dipastikan terus berkurang. Hal tersebut menciptakan minat akan properti, khususnya tempat tinggal, bakal tetap melambung tinggi. Pada waktu seseorang ingin memiliki unit properti muncul kendala yang umum dijumpai, yakni

sedikitnya data informasi terkait harga pasaran properti tersebut. Minimnya informasi terkait harga tanah dan bangunan pada akhirnya membuat banyak masyarakat melepas atau memiliki properti dengan tidak tahu informasi tentang harga pasaran. Jika konsumen memiliki properti dengan harga di atas pasaran, maka konsumen menghamburkan sumber daya lebih dari yang sebenarnya. Sumber daya yang dihamburkan umumnya tidak hanya berpatokan materi, tenaga dan waktu pun masuk ke jenis tersebut.

Pada masa modern, sebenarnya data terkait properti dapat dicari melalui internet dengan gampang. Data terkait umumnya meluas pada situs transaksi properti, situs resmi pemerintah, maupun dalam media digital. Tetapi, bukan seluruhnya data yang tersedia pada situs terkait bisa menggambarkan harga properti di wilayah tertentu. Dengan menerapkan inovasi *big data*, data-data terkait bisa saja diperoleh dan dikelompokkan dengan ukuran dan jumlah yang banyak berasal dari berbagai sumber (Suakanto, 2018).

## METODOLOGI PENELITIAN

Pada penelitian yang dilakukan, metodologi yang dipakai adalah pendekatan analisis regresi sederhana yaitu pendekatan yang dilakukan untuk mengetahui sebab akibat atau hubungan kesenjangan penjualan properti pada masing-masing wilayah. Dari data tersebut maka akan dilakukan analisis dan diperoleh rekomendasi terbaik yang dapat diberikan untuk meningkatkan penjualan properti.

### Pengumpulan Data

Untuk pengumpulan data telah disediakan dataset dari *website kaggle* dengan data penjualan properti pada masing-masing wilayah di NYC pada tahun 2016-2017. Data tersebut terdiri atas data letak wilayahnya, data tipe class properti, tipe bangunan, tingkatan pajak, dll.

Neighborhood	Building Class	Tax Class	Block	Lot
FLUSHING NORTH	01 ONE FAMILY D...	1	46%	
UPPER EAST SIDE L...	02 TWO FAMILY D...	2	37%	
Other (79744)	Other (50485)	Other (14998)	18%	16.3k
ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2A	392	6
ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2	399	26
ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2	399	39
ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2B	482	21
ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2A	484	55
ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2	485	16
ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2B	485	32
ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2	487	18

Gambar 3.1 Dataset mentah penjualan properti di NYC pada *Website Kaggle*

## HASIL DAN PEMBAHASAN

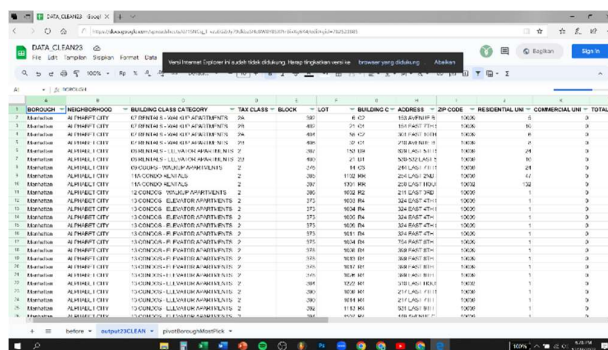
### A. Pengolahan Data

Setelah mendapatkan dataset yang diperlukan untuk melakukan analisis, selanjutnya adalah proses pengolahan data tersebut yakni:

#### 1. Melakukan Data Cleaning

Data cleaning maupun membersihkan data adalah tahap yang dipakai dengan tujuan melacak, memperbaiki ataupun menghilangkan dataset, tabel, dan database yang *error* atau kurang valid. definisi berikut berpatok pada data kotor yang akan diubah, dimodifikasi atau dihilangkan setelah proses pendefinisian data yang tidak lengkap, tidak benar, tidak tepat, dan tidak relevan. Tahap data cleaning ini sangat vital dengan tujuan mengantisipasi data berulang, menjalankan data agar lebih terstruktur, dan dapat berkaitan. Tahap data cleaning mencakup 5 proses yaitu; (1) Analisis Data, (2) mengartikan *workflow* transformasi dan peraturan pemetaan, (3) verifikasi, (4) Transformasi, (5) Pembalikan data yang sudah dibersihkan.

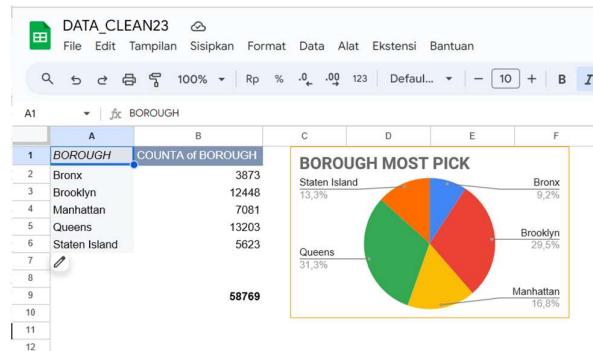
Langkah nomor satu dalam tahap Data cleaning yakni analisis data untuk identifikasi ketidak-akuratan dan ketidak-konsistenan yang terkait di *database*. Hasil nyayakni prediksi untuk setiap prediksi anomali apakah hal terkait terjadi di dalam *database*. Berikutnya, menerjemahkan *workflow* transformasi dan penghilangan anomali yang dijalankan berdasarkan urutan pelaksanaan pada data. Hal tersebut dipecah sesudah analisis data untuk memperoleh informasi tentang anomali terkait. Langkah ketiga yakni proses verifikasi. Dalam fase ini, ketepatan dan keefektifan transformasi alur kerja ditinjau ulang. Fase ini terdiri dari beberapa iterasi untuk mengecek semua ketidak-akuratan, Sesudah data ditinjau dan divalidasi, tahap-tahap perubahan akan dijalankan (Widiari, 2020).



BUILDING ID	BUILDING CLASS CATEGORY	TAX CLASS	BLOCK	LOT	ADDRESS	ZIP CODE	RESIDENTIAL UNITS	COMMERCIAL UNITS	TOTAL
1	13-00000	1	100	100	100	100	1	0	1
2	13-00000	1	100	100	100	100	1	0	1
3	13-00000	1	100	100	100	100	1	0	1
4	13-00000	1	100	100	100	100	1	0	1
5	13-00000	1	100	100	100	100	1	0	1
6	13-00000	1	100	100	100	100	1	0	1
7	13-00000	1	100	100	100	100	1	0	1
8	13-00000	1	100	100	100	100	1	0	1
9	13-00000	1	100	100	100	100	1	0	1
10	13-00000	1	100	100	100	100	1	0	1
11	13-00000	1	100	100	100	100	1	0	1
12	13-00000	1	100	100	100	100	1	0	1
13	13-00000	1	100	100	100	100	1	0	1
14	13-00000	1	100	100	100	100	1	0	1
15	13-00000	1	100	100	100	100	1	0	1
16	13-00000	1	100	100	100	100	1	0	1
17	13-00000	1	100	100	100	100	1	0	1
18	13-00000	1	100	100	100	100	1	0	1
19	13-00000	1	100	100	100	100	1	0	1
20	13-00000	1	100	100	100	100	1	0	1

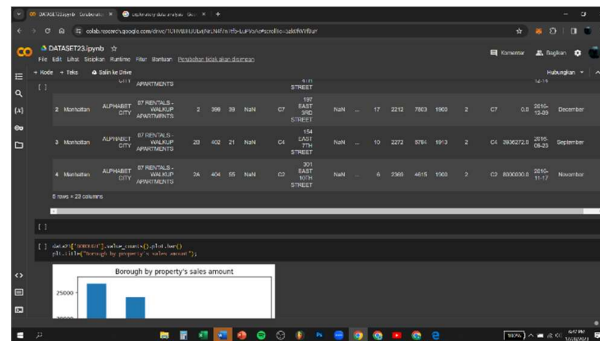
Gambar 3.2 Proses Data Cleaning pada Spreadsheet

Membuat pivot sederhana dengan membagi data berdasarkan wilayah yang tertera pada dataset meliputi Staten islan, Queens, Bronx, Brooklyn dan Manhattan.



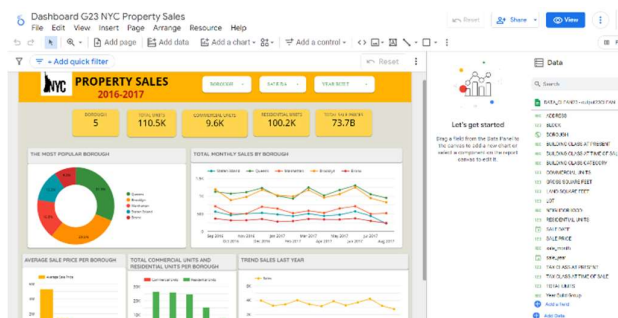
Gambar 3.3 Pivot penjualan properti di NYC pada *Spreadsheet*

Memproses data dengan Python untuk melakukan *exploratory data analysis* tujuannya untuk pendekatan analisis set data untuk meringkaskan ciri utamanya dengan menggunakan grafik statistik dan kaedah visualisasi data lain.



Gambar 3.4 *Exploratory Data analysis* dengan python

Membuat *dashboard* melalui website *Looker Studio* yang berfungsi untuk visualisasi data dengan grafik agar mudah untuk dipahami.



Gambar 3.5 *Dashboard* Data Penjualan properti di NYC

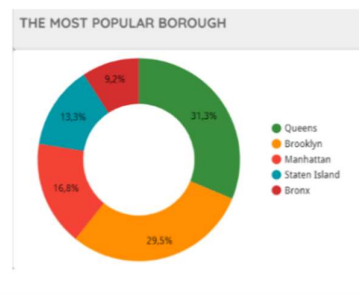
Membuat *deck* untuk proses *data communication*, agar *stakeholders* terkait dapat memahami analisis dari data yang tersedia untuk pengambilan keputusan bisnis.



Gambar 3.6 Deck Data Penjualan properti di NYC

Membuat analisis dan kesimpulan berdasarkan dataset dan visualisasi data yang telah dibuat sebelumnya, dengan analisis sebagai berikut :

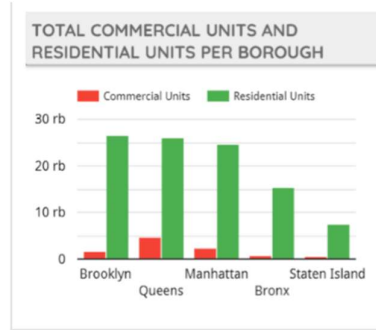
Queens adalah wilayah paling populer di Kota New York, dengan populasi 31,3%. Dan Bronx adalah wilayah yang paling tidak populer, dengan populasi 9,2%. Untuk mengatasi kesenjangan ini, kota ini dapat fokus pada inisiatif untuk meningkatkan kualitas hidup di Bronx.



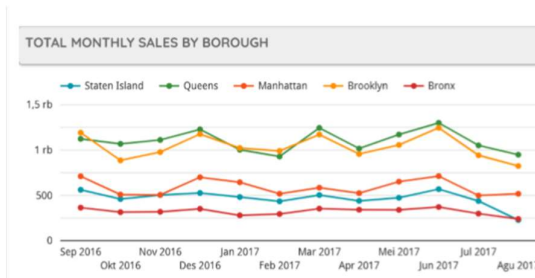
Gambar 3.7 Grafik perbandingan wilayah yang paling populer di NYC

Di seluruh wilayah, jumlah unit hunian jauh melebihi unit komersial. Brooklyn memiliki jumlah unit hunian tertinggi (26.581), sedangkan Staten Island memiliki jumlah unit komersial terendah (488). Fokus diperlukan pada inisiatif yang menarik dunia usaha ke Staten Island, sehingga menciptakan lingkungan ekonomi yang lebih seimbang dan berkelanjutan.





Gambar 3.8 Grafik perbandingan total unit komersial dan unit residential di NYC Queens memiliki penjualan bulanan tertinggi, dengan puncaknya sebesar 1,301 pada bulan Juni 2017. Staten Island memiliki penjualan bulanan terendah, dengan titik terendah sebesar 228 pada bulan Agustus 2017. Kesenjangan penjualan bulanan antara Queens dan Staten Island merupakan masalah serius yang perlu diatasi. Kota harus fokus pada inisiatif untuk meningkatkan lingkungan bisnis di Staten Island.

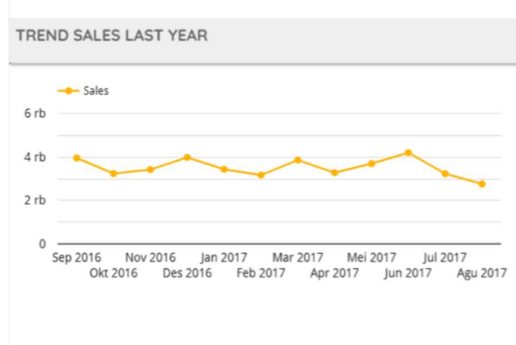


Gambar 3.9 Grafik perbandingan penjualan bulanan properti di NYC Manhattan memiliki harga jual rata-rata tertinggi, yaitu 5.333.841,66. (\$5 juta) Staten Island memiliki harga jual rata-rata terendah, dengan 518.950,35. (\$500 ribu) Kota harus fokus pada inisiatif untuk meningkatkan pasar perumahan di Staten Island. Tingginya harga jual rata-rata di Manhattan kemungkinan besar disebabkan oleh tingginya permintaan perumahan di kawasan tersebut, serta terbatasnya pasokan lahan.



Gambar 3.10 Rata-rata harga jual per wilayah di NYC

- Penjualan tertinggi terjadi pada Juni 2017 sebanyak 4.201 unit. Penjualan terendah terjadi pada Agustus 2017 sebanyak 2.760 unit. Tren penjualan tahun lalu memprihatinkan. Pemerintah kota harus mengambil tindakan untuk memastikan penjualan lebih stabil di masa depan.



Gambar 3.11 Tren Penjualan tahun terakhir

## KESIMPULAN DAN SARAN

### A.KESIMPULAN

Setelah dilakukan penjelasan paparan pengolahan data yang sudah dilaksanakan maka didapatkan kesimpulan dari dataset penjualan properti di NYC sebagai berikut:

1. Queens adalah wilayah terpopuler nomor 1 di NYC, sedangkan Bronx adalah wilayah paling tidak populer.
2. Jumlah Unit Perumahan jauh melebihi Unit Komersial di seluruh kota.
3. Brooklyn memimpin dalam Unit Perumahan, sementara Staten Island tertinggal.
4. Manhattan memiliki harga jual rata-rata tertinggi, sedangkan Staten Island memiliki harga jual terendah.
5. Penjualan mencapai puncaknya pada bulan Juni 2017 dan mencapai titik terendah pada bulan Agustus 2017.

### B.SARAN

Berdasarkan pemaparan pengolahan data yang telah dilaksanakan diatas maka didapatkan saran dan rekomendasi dari dataset penjualan properti di NYC agar penjualan properti dapat meningkat 10% sebagai berikut :

1. Untuk Wilayah Bronx, Sebaiknya infrastruktur diperbaiki dan lebih ditingkatkan lagi dan juga lebih mempromosikan keberagaman dan inklusivitas.

2. Untuk Wilayah Staten Island, Sebaiknya menarik Pengusaha untuk mau berinvestasi pada wilayah tersebut, meningkatkan Infrastruktur dalam wilayah, menyediakan perumahan yang terjangkau dan menstabilkan Penjualan.

#### **DAFTAR PUSTAKA**

- Luthfiana, M. (2020). ANALISIS JURNAL STATISTIKA DALAM PENGELOLAAN DATA DAN NILAI RAPORT SISWA DI SDN PANGKAH WETAN. *Statistics Journal Analysis In Data Management And Student Report Value At SDN Pangkah Wetan*.
- Rizal., Martanto & Yudhistira. A. (2022). ANALISA DATASET SOFTWARE DEFINED NETWORK INTRUSION MENGGUNAKAN ALGORITMA DEEP LEARNING H2O. *Jurnal Mahasiswa Teknik Informatika*.
- Suakanto, S., Aldi, C., Ventje, J. L & Dina, A. (2018). Pengembangan Sistem Prediksi Harga Pasar Properti Menggunakan Big Data Platform. *Jurnal Telematika, vol. 13 no. 1*
- Susanto, P & Gatot, I. K. (2023). Analisis Terbatasnya Peminatan Profesi Data Analyst Di Indonesia Berdasarkan Pendekatan Analytical Hierarchy Process (AHP). *Journal of Information System, Applied, Management, Accounting and Research*.
- Widiari, A. P., I Made, A. D., & Dwi, P. (2020). Teknik Data Cleaning Menggunakan Snowflake untuk Studi Kasus Objek Pariwisata di Bali. *JURNAL ILMIAH MERPATI VOL. 8, NO. 2*